MULTIPARAMETER ERGODIC AVERAGES FOR TWO COMMUTING ACTIONS OF AN AMENABLE GROUP

JOHN T. GRIESMER

Abstract. We find limits of some multiple ergodic averages, generalizing a result of Bergelson to the setting of two commuting transformations and actions of amenable groups.

1. Introduction

In order to give an ergodic-theoretic proof of Szemerédi's theorem on arithmetic progressions, Furstenberg ([9]) analyzed the behavior of the averages

$$\frac{1}{N-M}\sum_{n=M}^{N-1}\int f\cdot f\circ T^n\cdot f\circ T^{2n}\cdot\cdots\cdot f\circ T^{kn}\,d\mu, \tag{1}$$

where $T : X \to X$ is a transformation of a measure space $(X, \mathcal{B}, \mu)$ which preserves a probability measure μ, and $f \in L^\infty(\mu)$. This raised the problem of deciding whether such averages converge as $N - M \to \infty$, and if they do converge, identifying the limit. The case $k = 2$ was settled affirmatively in [9], and the case $k = 3$ was settled affirmatively by Conze and Lesigne in [6] and [7]. Zhang [20] generalized the results of [7] to the case where T^n, T^{2n}, and T^{3n} are replaced by T^n, S^n, and R^n, where R, S, and T are commuting, ergodic transformations, and the transformations $R^{-1}S, S^{-1}T$, and $T^{-1}R$ are assumed to be ergodic. Finally, the question of convergence in (1) was settled affirmatively for all k by Host and Kra in [15], and independently by Ziegler in [19]. Each of the aforementioned papers uses a method of reducing the problem to studying the appropriate *characteristic factors*, a notion which we explain in the next section. In these proofs, identifying the appropriate characteristic factors is an essential step in identifying the limit. Recently, Tao has generalized this convergence result to the case where the powers of T in (1) are replaced by commuting transformations $T_1, \dots, T_k$. Interestingly, Tao's proof ([18]) does not use the method of characteristic factors, and does not identify the limit of the averages. T. Austin ([1]) has found another, more general proof of Tao's theorem which does not identify the characteristic factors. The proof in [1] uses properties of diagonal measures, as in the proof of Theorem 2.3 and Corollary 2.4 of [9].

The following related result is due to Bergelson, and appears as Theorem 5.3 in [3]. Here a *syndetic* subset of $\mathbb{Z}^2$ is one that meets every square of side length L for some $L > 0$.

Theorem 1.1. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible probability measure preserving system. Then:*

(1) *For any $f_1, f_2, f_3 \in L^\infty(X, \mathcal{B}, \mu)$,*

$$\lim_{N-M\to\infty}\frac{1}{(N-M)^2}\sum_{n,m=M}^{N-1} f_1(T^n x)f_2(T^m x)f_3(T^{n+m}x)$$

Date: December 10, 2008.

2000 *Mathematics Subject Classification.* 37A05.

exists in $L^2(\mu)$.

(2) *For any* $A \in \mathcal{B}$ *with* $\mu(A) > 0$

$$\lim_{N-M\to\infty} \frac{1}{(N-M)^2} \sum_{n,m=M}^{N-1} \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) \geq \mu(A)^4.$$

From this fact, Bergelson derives the following combinatorial corollary which generalizes Khintchine's refinement of the Poincaré recurrence theorem.

Corollary 1.2. *For* $(X, \mathcal{B}, \mu, T)$ *as in Theorem 1.1, for any* $0 < \lambda < 1$ *and any* $A \in \mathcal{B}$ *with* $\mu(A) > 0$, *the set*

$$\{(n,m) \in \mathbb{Z}^2 : \mu\left((A \cap T^n A) \cap T^m(A \cap T^n A)\right) > \lambda \mu(A)^4\}$$

is syndetic.

Using Furstenberg's correspondence principle (see [2]) one can derive a combinatorial consequence of Corollary 1.2, which says roughly that given a set of positive density δ in $\mathbb{Z}$, there is a syndetic set of $(n,m) \in \mathbb{Z}^2$ so that for all such (n,m), configurations of the form $(a, a+n, a+m, a+n+m)$ appear as frequently as you would expect in a random set having density δ. The precise statement follows.

Theorem 1.3. *Let* $A \subseteq \mathbb{Z}$ *such that* $d^*(A) = \limsup_{N-M\to\infty} \frac{|A\cap[M,N-1]|}{N-M} = \delta$. *Then for all* $\varepsilon > 0$, *the set of* (n,m) *such that*

$$d^*(A \cap (A-n) \cap (A-m) \cap (A-n-m)) > \delta^4 - \varepsilon$$

is syndetic.

It is natural to conjecture that a similar result holds for sets of positive density in $\mathbb{Z}^2$, and that such a result may be proved by establishing convergence for averages of the form

$$\frac{1}{(N_1-M_1)(N_2-M_2)} \sum_{n=M_1}^{N_1-1} \sum_{m=M_2}^{N_2-1} f_1(x) f_2(T^n x) f_3(S^m x) f_4(T^n S^m x), \tag{2}$$

where T and S are commuting, measure preserving transformations of a probability measure space. In this paper we establish convergence of the averages in (2) and use this to deduce a combinatorial consequence for sets of positive density in $\mathbf{Z}^2$.

We actually consider a more general situation, which may be motivated by studying configurations in sets of positive density in an amenable group G, and in the cartesian square $G \times G$. We say a group is *amenable* if there is a sequence of finite sets $\Phi_n \subseteq G$ such that for all $g \in G, \lim_{n\to\infty} \frac{|\Phi_n \cap g\Phi_n|}{|\Phi_n|} = 1$. Such a sequence is called a *left Følner sequence*, and these sequences allow one to define a translation-invariant notion of density in the group G. Analogously, we say that Φ_n is a *right* Følner sequence if for all $g \in G, \lim_{n\to\infty} \frac{|\Phi_n \cap \Phi_n g|}{|\Phi_n|} = 1$, and we say that Φ_n is a *two-sided* Følner sequence if it is both a left- and a right Følner sequence.

Supposing that G is an amenable group and that $g \mapsto S_g, g \mapsto T_g$ are commuting actions of G by measure preserving transformations on a standard probability space $(X, \mathcal{B}, \mu)$, we demonstrate that the limit

$$\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{g\in\Phi_n, h\in\Psi_n} f_1(T_g x) f_2(S_h x) f_3(T_g S_h x)$$

exists for all two-sided Følner sequences Φ, Ψ in G and that the limit is independent of the choice of Følner sequences. We identify the limit, and also give a combinatorial inequality analogous to Corollary 1.2. Further, following [5], we give necessary and sufficient conditions for the limit to be equal to a constant almost everywhere.

Our main result is the following theorem.

Theorem 1.4. *Let G be a countable amenable group, let $\mathbf{X} = (X, \mathcal{B}, \mu)$ be a standard probability space with probability measure μ, and commuting measure preserving actions T, S of G on X. Then*

(1) For all $f_1, f_2, f_3 \in L^\infty(\mu)$, and all two-sided Følner sequences Φ, Ψ in G, the limit

$$L = \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} f_1(T_g x) f_2(S_h x) f_3(T_g S_h x) \tag{3}$$

exists in $L^2(\mu)$.

(2) L is equal to a constant μ-almost everywhere for all $f_1, f_2, f_3 \in L^\infty(\mu)$ if and only if $T \times T$ and $S \times S$ are ergodic.

The condition that $T \times T$ is ergodic is of course equivalent to the condition that T is weakly mixing.

It is not much of a sacrifice to restrict our attention to two-sided Følner sequences, for every countable amenable group admits two-sied Følner sequences - see [13] for a proof.

Theorem 1.4 has a corollary analogous to Corollary 1.2.

Corollary 1.5. *Let $(X, \mathcal{B}, \mu), T$ and S be as in Theorem 1.4. Suppose $f \in L^\infty(\mu)$ is a nonnegative function. Then for all two-sided Følner sequences Φ, Ψ*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} \int f T_g f S_h f T_g S_h f \, d\mu \geq \left(\int f \, d\mu \right)^4. \tag{4}$$

We derive several consequences from Corollary 1.5. First there is the natural combinatorial result about sets of positive density in $G \times G$. In a general group G, we say a set $S \subseteq G$ is *left syndetic* if there is a finite set $F \subseteq G$ such that $G = \bigcup_{g\in F} gS$, and we say that S is *right* syndetic if there is a finite set F such that $G = \bigcup_{g\in F} Sg$.

Corollary 1.6. *Let $E \subseteq G\times G$, and let Φ be a Følner sequence in $G\times G$ with $\limsup_{n\to\infty} \frac{|E\cap\Phi_n|}{|\Phi_n|} = \delta > 0$. Then for all $\varepsilon > 0$, the set*

$$\left\{ (g,h) : \limsup_{n\to\infty} \frac{|E \cap E(g, 1_G) \cap E(1_G, h) \cap E(g,h) \cap \Phi_n|}{|\Phi_n|} > \delta^4 - \varepsilon \right\}$$

is both left- and right syndetic in $G \times G$

A combinatorial corollary about partitions of subsets in $G \times G \times G$ will also be derived in Section 5.

In Section 6 we specialize to the case $G = \mathbb{Z}$, and use the main theorem to construct a class of examples which illustrates the difference between the case of powers of a single transformation as studied in [15] and the case of commuting transformations.

2. Acknowledgements.

The author thanks Vitaly Bergelson for posing the problem of establishing Theorem 1.4 and for helpful comments and advice. Thanks are due Alexander Leibman for suggesting the example in Section 6.

3. Notation and background.

Throughout the paper, we fix a probability space $(X, \mathcal{B}, \mu)$. We may assume, without loss of generality, that $(X, \mathcal{B}, \mu)$ is a Lebesgue space - that is, measure-theoretically isomorphic to the unit interval with Lebesgue measure and the completed Borel σ-algebra, together with possibly countably many atoms. (See [17], Chapter 15, Theorem 4, and Chapter 15, Theorem 20, for justification.) In particular we can assume that X is compact metric.

As mentioned above, every discrete amenable group has a two-sided Følner sequence. We write Φ for a Følner sequence $\{\Phi_n\}_{n\in\mathbb{N}}$. If Φ is a Følner sequence and $g \in G$, we write $g\Phi$ and Φg for the sequences $\{g\Phi_n\}, \{\Phi_n g\}$ respectively, and note without proof that if Φ is a (left-, right-, two-sided) Følner sequence then $g\Phi$ and Φg are both (left-, right-, two-sided) Følner sequences.

We denote by T or S an action of an amenable discrete group G on $(X, \mathcal{B}, \mu)$, and write T_g for the transformation corresponding to g, so that $T_g T_h = T_{gh}$ for all $g, h \in G$. Note that this induces an anti-action of G on $L^2(\mu)$ by $T_g f = f \circ T_g$ for $f \in L^2(\mu)$. We will use $T_g f$ to denote $f \circ T_g$. Write $T \times T$ for the action of G on $X \times X$ given by $g \mapsto T_g \times T_g$.

We briefly review the notions of a factor of a dynamical system, conditional expectation, and disintegration of measures. Write $\mathbf{X} = (X, \mathcal{B}, \mu, T, S)$ for a dynamical system where $(X, \mathcal{B}, \mu)$ is as above, and T and S are commuting actions of G on $(X, \mathcal{B}, \mu)$ by measure preserving transformations of $(X, \mathcal{B}, \mu)$. We say that $\mathbf{Y} = (Y, \mathcal{D}, \nu, T', S')$ is a *factor* of $\mathbf{X}$ if there exists $\pi : X \to Y$ with $\pi^{-1}(\mathcal{D}) \subseteq \mathcal{B}, \mu(\pi^{-1}(A)) = \nu(A)$ for all $A \in \mathcal{D}$, and $T'_g \pi(x) = \pi(T_g x)$ and $S'_g \pi(x) = \pi(S_g x)$ for μ-almost all x. One can identify $\mathcal{D}$ with the σ-algebra $\pi^{-1}(\mathcal{D})$.

Given a T- and S-invariant, countably generated sub σ-algebra $\mathcal{B}' \subset \mathcal{B}$, there is always be a corresponding factor $(Y, \mathcal{D}, \nu, T', S')$ so that $\mathcal{B}' = \{\pi^{-1}(D) : D \in \mathcal{D}\}$ (up to sets of measure 0), and we will abuse notation and identify $\mathcal{B}'$ and $\mathcal{D}$.

If $\mathcal{A}$ is a sub σ-algebra of $\mathcal{B}$, and $f \in L^2(\mu)$, one can define the conditional expectation (relative to μ) $\mathbb{E}(f|\mathcal{A})$ by $\mathbb{E}(f|\mathcal{A})(x) := Pf(x)$, where Pf is the orthogonal projection (in $L^2(\mu)$) of f onto the subspace of $\mathcal{A}$-measurable functions. (This can be extended to a map on $L^1(\mu)$, but we have no need for this here.)

The conditional expectation can be used to define disintegration of measures, as in [10], Theorem 5.8. If $\mathcal{A}$ is a countably generated sub σ-algebra of $\mathcal{B}$, then there is a measurable map $x \mapsto \mu_x$ from X to the space of regular Borel measures on X (with the weak* topology), such that

- i For every $f \in L^1(\mu), f \in L^1(X, \mathcal{B}, \mu_x)$ for μ-almost every x, and $\mathbb{E}(f|\mathcal{A})(x) = \int f \, d\mu_x$ for μ-almost every x.
- ii $\int \int f \, d\mu_x \, d\mu(x) = \int f \, d\mu$ for every $f \in L^1(X, \mathcal{B}, \mu)$.

If $(Y, \mathcal{D}, \nu, T', S')$ is a factor of $\mathbf{X}$ with factor map π, we use $x \mapsto \mu_{\pi(x)}$ to denote the disintegration of μ over $\pi^{-1}(\mathcal{D})$. Note that $\int f d\mu_{\pi(T_g x)} = \int f \circ T_g \, d\mu_{\pi(x)}$ for μ-almost every x.

Let $\mu = \int \mu_{\pi(x)}\, d\mu(x)$ be the disintegration of μ over $\mathcal{D}$, so that $x \mapsto \mu_{\pi(x)}$ is a $\mathcal{D}$-measurable function, and $\int f\, d\mu_x = \mathbb{E}(f|\mathcal{D})$ for μ-almost every x. One defines the *relative product* measure $\mu \underset{\mathcal{D}}{\times} \mu$ on $(X \times X, \mathcal{B} \otimes \mathcal{B})$ by $\mu \underset{\mathcal{D}}{\times} \mu = \int \mu_{\pi(x)} \times \mu_{\pi(x)}\, d\mu(x)$. This means that $\int f\, d\mu \underset{\mathcal{D}}{\times} \mu = \int \int f\, d\mu_{\pi(x)} \times \mu_{\pi(x)}\, d\mu(x)$ for bounded $\mathcal{B} \otimes \mathcal{B}$-measurable f. Equivalently, one may define $\mu \underset{\mathcal{D}}{\times} \mu$ by $\int f \otimes g d\mu \underset{\mathcal{D}}{\times} \mu = \int \mathbb{E}(f|\mathcal{D})\mathbb{E}(g|\mathcal{D})\, d\mu$ for bounded f, g. The relative product system $\mathbf{X} \underset{\mathbf{Y}}{\times} \mathbf{X}$ is defined as the system $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \underset{\mathcal{D}}{\times} \mu, T \times T, S \times S)$.

If R is a measure preserving action of G on $(X, \mathcal{B}, \mu)$, define $\mathcal{I}_R$ to be the σ-algebra of R-invariant sets (up to sets of measure 0): $\mathcal{I}_R = \{A \in \mathcal{B} : \mu(R_g A \triangle A) = 0 \text{ for all } g \in G\}$. Let $\mathbf{I}_R = (I_R, \mathcal{I}_R, \mu, G)$ be the factor of $\mathbf{X}$ corresponding to $\mathcal{I}_R$. We will mainly be concerned with $\mathbf{I}_T$ and $\mathbf{I}_S$. Let $\tau : X \to I_T$ and $\sigma : X \to I_S$ denote the respective factor maps.

We will use a version of the ergodic theorem for amenable groups.

Theorem 3.1. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be a measure preserving system with T a measure preserving action of G on X. Then for all two-sided Følner sequences Φ in G, and all $f \in L^2(\mu)$, the expectation $\mathbb{E}(f|\mathcal{I}_T)$ is given by*

$$\mathbb{E}(f|\mathcal{I}_T) = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} f \circ T_g = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} f \circ T_g^{-1}$$

where the convergence is in the sense of $L^2(\mu)$.

4. Proof of the main theorem.

We restate our main theorem for convenience.

Theorem 1.4. *Let G be a countable amenable group, let $\mathbf{X} = (X, \mathcal{B}, \mu)$ be a standard probability space with probability measure μ, and commuting measure preserving actions T, S of G on X. Then*

(1) For all $f_1, f_2, f_3 \in L^\infty(\mu)$, and all two-sided Følner sequences Φ, Ψ in G, the limit

$$L = \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} f_1(T_g x) f_2(S_h x) f_3(T_g S_h x) \tag{5}$$

exists in $L^2(\mu)$.

(2) L is equal to a constant μ-almost everywhere for all $f_1, f_2, f_3 \in L^\infty(\mu)$ if and only if $T \times T$ and $S \times S$ are ergodic.

Our proof will follow the proof of Theorem 1.1 found in [3], which is an example of the method of characteristic factors. See [11],[12], or [15] for a general description of this method. Briefly, if $(X, \mathcal{B}, \mu, T)$ is a measure preserving system as in Theorem 1.1, we say that a factor $\mathbf{Y} = (Y, \mathcal{D}, \nu, T')$ of a $\mathbf{X}$ is a *characteristic factor for the scheme* (T^n, T^m, T^{n+m}) if for all $f_1, f_2, f_3 \in L^\infty(\mu)$,

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} T^n f_1 S^m f_2 T^{n+m} f_3 - T^n \mathbb{E}(f_1|\mathcal{D}) T^m \mathbb{E}(f_2|\mathcal{D}) T^{n+m} \mathbb{E}(f_3|\mathcal{D}) = 0.$$

Thus, if $\mathbf{Y}$ is a characteristic factor for the scheme (T^n, T^m, T^{n+m}) one need only establish the convergence result for functions defined on $\mathbf{Y}$. In [3] it is shown that the Kronecker factor is a characteristic factor for the scheme (T^n, T^m, T^{n+m}) and convergence is established using

properties of eigenfunctions. For the sake of brevity, we define the notion of characteristic factor precisely only for our scenario.

Definition. Let $\mathbf{X}$, T and S be as in the statement of Theorem 1.4. We say that a factor $\mathbf{Y}$ of $\mathbf{X}$ is a *characteristic factor for the scheme* (T_g, S_h, T_gS_h) if for all functions $f_1, f_2, f_3 \in L^\infty(\mu)$ and two-sided Følner sequences Φ, Ψ,

$$\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_gf_1S_hf_2T_gS_hf_3 - T_g\mathbb{E}(f_1|\mathcal{D})S_h\mathbb{E}(f_2|\mathcal{D})T_gS_h\mathbb{E}(f_3|\mathcal{D}) = 0$$

Using similar methods to [4], we will find a characteristic factor for the scheme (T_g, S_h, T_gS_h). However, in the course of the proof, we only use *partial* characteristic factors: we find factors $\mathbf{Y}_1, \mathbf{Y}_2$ with corresponding invariant σ-algebras $\mathcal{D}_1, \mathcal{D}_2 \subseteq \mathcal{B}$ so that

$$\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_gf_1S_hf_2T_gS_hf_3 - T_g\mathbb{E}(f_1|\mathcal{D}_1)S_h\mathbb{E}(f_2|\mathcal{D}_2)T_gS_hf_3 = 0.$$

As with most applications of characteristic factors, we need a van der Corput lemma. The following appears as Lemma 4.2 in [4].

Lemma 4.1. *Suppose that $\{u_g : g \in G\}$ is a bounded set in a Hilbert space H, and that Φ is a left Følner sequence in G. If*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|^2} \left(\limsup_{m\to\infty} \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \sum_{h,k\in\Phi_n} \langle u_{hg}, u_{kg}\rangle \right) = 0,$$

then $\lim_{n\to\infty} \|\frac{1}{|\Phi_n|}\sum_{g\in\Phi_n} u_g\| = 0.$

We will actually apply a corollary of Lemma 4.1.

Corollary 4.2. *Suppose that $\{u_{g,h} : g, h \in G\}$ is a bounded set in a Hilbert space H, and that Φ, Ψ are two-sided Følner sequences in G. If*

$$\lim_{m\to\infty} \frac{1}{|\Phi_m|^2|\Psi_m|^2} \left(\limsup_{m\to\infty} \frac{1}{|\Phi_m||\Psi_m|} \sum_{\substack{j,j'\in\Phi_m\\ k,k'\in\Phi_m}} \sum_{\substack{g\in\Phi_n\\ h\in\Psi_n}} \langle u_{j'g,k'h}, u_{jg,kh}\rangle \right) = 0$$

then $\lim_{n\to\infty} \|\frac{1}{|\Phi_n||\Psi_n|}\sum_{g,h\in\Phi_n\times\Psi_n} u_{g,h}\| = 0.$

In order to describe our partial characteristic factors, we use the construction of the maximal isometric extension of a factor of $\mathbf{X}$. If $\mathbf{Y}$ is a factor of $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ let $W_\mathbf{Y}$ denote the closed T-invariant subspace of $L^2(\mu)$ spanned by functions of the form $k(x) = \int H(x,z)\phi(z)\, d\mu_{\pi(x)}(z)$, where $H \in L^\infty(\mu \underset{\mathcal{D}}{\times} \mu)$ is $T\times T$-invariant (with respect to $\mu \underset{\mathcal{D}}{\times} \mu$), and $\phi \in L^\infty(\mathbf{X})$. In fact, $W_\mathbf{Y}$ is a closed algebra of functions, and therefore determines a factor of $\mathbf{X}$, but this fact will not be needed in our proofs. To check that $W_\mathbf{Y}$ is T-invariant, we write

$$\begin{aligned} \int H(Tx,z)\phi(z)\, d\mu_{\pi(Tx)}(z) &= \int H(Tx,Tz)\phi(Tz)\, d\mu_{\pi(x)}(z) \\ &= \int H(x,z)\phi(Tz)\, d\mu_{\pi(x)}(z). \end{aligned} \tag{6}$$

The next lemma is essentially Lemma 7.6 from [10], adapted to the setting of amenable groups.

Lemma 4.3. *Let* $\mathbf{Y}$ *be a factor of* $\mathbf{X}$*. If* $f \in L^\infty(\mathbf{X})$ *and* $f \perp W_{\mathbf{Y}}$*, then for all two-sided Følner sequences* Φ*,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \|\mathbb{E}(fT_g f|\mathcal{D})\| = 0.$$

Proof. Let Φ be a two-sided Følner sequence. It suffices to show that

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \|\mathbb{E}(fT_g f|\mathcal{D})\|^2 = 0.$$

Write $\|E(fT_g f|\mathcal{D})\|^2 = \int \int fT_g f \, d\mu_{\pi(x)} \int \bar{f} T_g \bar{f} \, d\mu_{\pi(x)} \, d\mu(x)$, so that

$$\|\mathbb{E}(fT_g f|\mathcal{D})\|^2 = \int \int f(w) f(T_g w) \bar{f}(T_g z) \bar{f}(z) \, d\mu_{\pi(x)} \times \mu_{\pi(x)}(w,z) \, d\mu(x). \tag{7}$$

Averaging over Φ in (7) yields

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \|\mathbb{E}(fT_g f|\mathcal{D})\|^2 = \int \int f(w) H(w,z) \bar{f}(z) \, d\mu_{\pi(x)} \times \mu_{\pi(x)}(w,z) \, d\mu(x), \tag{8}$$

where H is the projection of $f \otimes \bar{f}$ on the space of $T \times T$-invariant functions on $\mathbf{X} \underset{\mathbf{Y}}{\times} \mathbf{X}$. Changing the order of integration, we rewrite (8) as

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} & \|\mathbb{E}(fT_g f|\mathcal{D})\| \\ &= \int \int f(w) \int H(w,z) \bar{f}(z) \, d\mu_{\pi(x)}(z) \, d\mu_{\pi(x)}(w) \, d\mu(x) \\ &= \int f(x) \int H(x,z) \bar{f}(z) \, d\mu_{\pi(x)}(z) \, d\mu(x), \end{aligned} \tag{9}$$

where the last equality is justified by the disintegration $\mu = \int \mu_{\pi(x)} \, d\mu(x)$. By assumption, f is orthogonal to the function given by the inner integral $\int H(x,z) \bar{f}(z) \, d\mu_x(w)$, so the last integral in (9) is 0. □

4.1. **Reduction to partial characteristic factors.** Let $W_{T/S}$ denote the closed subspace of $L^2(\mu)$ spanned by functions of the form $k(x) = \int H(x,z)\phi(z) \, d\mu_{\sigma(x)}(z)$, where H is $T \times T$-invariant (with respect to $\mu \underset{\mathcal{I}_S}{\times} \mu$,) and $\phi \in L^\infty(\mu)$. Let $W_{S/T}$ denote the closed subspace of $L^2(\mu)$ spanned by functions of the form $k(x) = \int H(x,z)\phi(z) \, d\mu_{\tau(x)}(z)$, where H is $S \times S$-invariant (with respect to $\mu \underset{\mathcal{I}_S}{\times} \mu$). Let $P : L^2(\mu) \to W_{T/S}$ denote the orthogonal projection onto $W_{T/S}$, and let $Q : L^2(\mu) \to W_{S/T}$ denote the projection onto $W_{S/T}$.

The spaces $W_{T/S}$ and $W_{S/T}$ will serve as partial characteristic factors, in that we will reduce the problem of showing that the limit exists in (3) to the case where $f_1 \in W_{T/S}$ and $f_2 \in W_{S/T}$. We will not, however, show that these spaces are closed under pointwise multiplication, so we are not considering them as factors at present. Once convergence in (3) is established, we will see that $W_{T/S}$ and $W_{S/T}$ are closed algebras of functions, and are equal.

Let Φ and Ψ be two-sided Følner sequences in G. Let $f_1, f_2, f_3 \in L^\infty(\mu)$, and define a sequence

$$A_n := \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_g f_1 S_h f_2 T_g S_h f_3.$$

and auxiliary sequences

$$B_n^{(1)} := \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_g(Pf_1) S_h f_2 T_g S_h f_3$$
$$B_n^{(2)} := \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_g f_1 S_h(Qf_2) T_g S_h f_3$$
$$C_n := \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_g(Pf_1) S_h(Qf_2) T_g S_h f_3.$$

Lemma 4.4. *For $i = 1, 2$, we have $\lim_{n\to\infty} A_n - B_n^{(i)} = 0$, and $\lim_{n\to\infty} A_n - C_n = 0$.*

Proof. We prove only that $\lim_{n\to\infty} A_n - B_n^{(1)} = 0$. The case $i = 2$ is similar, and the last assertion follows from the first two.

Without loss of generality, suppose that $Pf_1 = 0$. We need to show that $\lim_{n\to\infty} A_n = 0$, so we apply Lemma 4.2. For $g, h \in G$ let $u_{g,h} = T_g f_1 S_h f_2 T_g S_h f_3$. Then for $g, j, j', h, k, k' \in G$,

$$\langle u_{j'g,k'h}, u_{jg,kh}\rangle = \int T_g(T_{j'} f_1 T_j \bar{f}_1) S_h(S_{k'} f_2 S_k \bar{f}_2) T_g S_h(T_{j'} S_{k'} f_3 T_j S_k \bar{f}_3)\, d\mu. \tag{10}$$

Applying $T_g^{-1} S_h^{-1}$ to the integrand in (10) yields

$$\langle u_{j'g,k'h}, u_{jg,kh}\rangle = \int S_h^{-1}(T_{j'} f_1 T_j \bar{f}_1) T_g^{-1}(S_{k'} f_2 S_k \bar{f}_2)(T_{j'} S_{k'} f_3 T_j S_k \bar{f}_3)\, d\mu. \tag{11}$$

Averaging over $\Phi \times \Psi$ in the variables g and h we get

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{|\Phi_n \times \Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} \langle u_{j'g,k'h}, u_{jg,kh}\rangle \\ &\qquad = \int \mathbb{E}(T_{j'} f_1 T_j \bar{f}_1 | \mathcal{I}_S) \mathbb{E}(S_{k'} f_2 S_k \bar{f}_2 | \mathcal{I}_T) T_{j'} S_{k'} f_3 T_j S_k \bar{f}_3\, d\mu. \end{aligned} \tag{12}$$

Applying the Cauchy-Schwartz inequality to the right-hand side of (12), and noting that $\|\mathbb{E}(S_{k'} f_2 S_k \bar{f}_2 | \mathcal{I}_T) T_{j'} S_{k'} f_3 T_j S_k \bar{f}_3\|$ is bounded by some nonnegative constant α, we have the inequality

$$\left| \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} \langle u_{j'g,k'h}, u_{jg,kh}\rangle \right| \le \alpha \|\mathbb{E}(T_{j'} f_1 T_j \bar{f}_1 | \mathcal{I}_s)\|. \tag{13}$$

Averaging over $j, j' \in \Phi_m$ on both sides of (13) and applying Lemma (4.3), we get

$$\lim_{m\to\infty} \frac{1}{|\Phi_m|^2|\Psi_m|^2} \sum_{j,j'\in\Phi_m} \sum_{k,k'\in\Psi_m} \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} \langle u_{j'g,k'h}, u_{jg,kh}\rangle = 0.$$

By Lemma 4.2, we have $\lim_{n\to\infty} A_n = 0$. □

4.2. **Construction of an invariant measure on $X \times X \times X$.** Lemma 4.4 allows us to assume that $f_1 \in W_{T/S}$ and $f_2 \in W_{S/T}$ when we consider limits of the averages in (3). In this subsection, we show that this assumption allows us to approximate the summands $T_g f_1 S_h f_2 T_g S_h f_3$ in (3) by the linear combinations of continuous linear images of sequences the form $(T_g \times \mathrm{id} \times T_g)(\mathrm{id} \times S_h \times S_h)(\phi_1 \otimes \phi_2 \otimes \phi_3)$, where $\phi_1 \otimes \phi_2 \otimes \phi_3$ is defined on $X \times X \times X$ with a $T \times \mathrm{id} \times T$- and $\mathrm{id} \times S \times S$-invariant Borel measure. To be precise, there is a measure space $(Y, \mathcal{D}, \nu)$, measure preserving actions T', S' of G on Y, and functions $\phi_1, \phi_2, \phi_3 \in L^\infty(\nu)$, and a continuous linear map $M : L^2(\nu) \to L^2(\mu)$ so that the sequence $u_{g,h} := T_g f_1 S_h f_2 T_g S_h f_3$ can be approximated by a linear combination of sequences of the form $v_{g,h} = M(T'_g \phi_1 S'_h \phi_2 T'_g S'_h \phi_3)$.

For each $x \in X$, let $\lambda_x = \mu_{\sigma(x)} \times \mu_{\tau(x)} \times \delta_x$, where δ_x is the measure defined by $\delta_x(A) = 1$ if and only if $x \in A$. Let $\lambda = \int \lambda_x \, d\mu(x)$, so that λ is a measure on $X \times X \times X$. We may also define λ by $\int f_1 \otimes f_2 \otimes f_3 \, d\lambda = \int \mathbb{E}(f_1|\mathcal{I}_S)\mathbb{E}(f_2|\mathcal{I}_T) f_3 \, d\mu$. This second definition makes it clear that λ is $T \times \mathrm{id} \times T$-invariant and $\mathrm{id} \times S \times S$-invariant.

We make the following standard observation about λ:

Observation 4.5. *Let $H \in L^\infty(\lambda), f \in L^\infty(\lambda)$, and let $\{f_n\}_{n\in\mathbb{N}}$ be a uniformly bounded sequence of functions in $L^\infty(\lambda)$ such that $\lim_{n\to\infty} \|f_n - f\|_{L^2(\lambda)} = 0$. Let $F(x) = \int H f \, d\lambda_x$ and for each n, let $F_n(x) = \int H f_n \, d\lambda_x$, so that $F_n \in L^\infty(\mu)$. Then $\lim_{n\to\infty} \|F_n - F\|_{L^2(\mu)} = 0$.*

We can now prove part (1) Theorem 1.4. *Proof of Theorem 1.4, part (1).* Our goal is to show that the limit in (3) exists. By Lemma 4.4, we may assume that $f_1 \in W_{T/S}$ and $f_2 \in W_{S/T}$. Thus, f_1 may be approximated in $L^2(\mu)$ by linear combinations of functions of the form $\int H_1(x, z_1)\phi_1(z_1) \, d\mu_{\sigma(x)}(z_1)$, where $H_1 \in L^\infty(\mathbf{X} \underset{\mathbf{I}_S}{\times} \mathbf{X})$ is $T \times T$-invariant and $\phi_1 \in L^\infty(\mu)$, while f_2 may be approximated in $L^2(\mu)$ by linear combinations of functions of the form $\int H_2(x, z_2)\phi_2(z_2) \, d\mu_{\tau(x)}(z_2)$, where $H_2 \in L^\infty(\mathbf{X} \underset{\mathbf{I}_T}{\times} \mathbf{X})$ is $S \times S$-invariant. Since each f_i is uniformly bounded, and each of the approximations to the f_i are uniformly bounded, we may deduce the existence of the limit in (3) from the existence of limits of the form

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} &\left(\int H_1(T_g x, z_1)\phi_1(z_1) \, d\mu_{\sigma(T_g x)}(z_1)\right) \times \\ &\left(\int H_2(S_h x, z_2)\phi_2(z_2) \, d\mu_{\tau(S_h x)}(z_2)\right) f_3(T_g S_h x).\end{aligned} \tag{14}$$

Write $u_{g,h}$ for the summand in (14). Computing as in equation (6), we can rewrite $u_{g,h}$ as the linear image of a function in $L^\infty(\lambda)$:

$$\begin{aligned} u_{g,h} &= \int H_1(T_g x, z_1)\phi_1(z_1) \, d\mu_{\sigma(T_g x)}(z_1) \int H_2(S_h x, z_2)\phi_2(z_2) \, d\mu_{\tau(S_h x)} f_3(T_g S_h x) \\ &= \int H_1(x, z_1)\phi_1(T_g z_1) \, d\mu_{\sigma(x)}(z_1) \int H_2(x, z_2)\phi_2(S_h z_2) \, d\mu_{\tau(x)} f_3(T_g S_h x) \\ &= \int H_1(x, z_1)H_2(x, z_2)\phi_1(T_g z_1)\phi_2(S_h z_2) f_3(T_g S_h z_3) \, d\mu_{\sigma(x)} \times \mu_{\tau(x)} \times \delta_x(z_1, z_2, z_3). \end{aligned} \tag{15}$$

Define a function $K \in L^\infty(\lambda)$ by $K(z_1, z_2, z_3) = H(x, z_1)H(x, z_2)$ if $\sigma(z_1) = \sigma(x)$ and $\tau(z_2) = \tau(x)$. From the last line of (15) we have

$$u_{g,h}(x) = \int K \cdot (T \times \mathrm{id} \times T)_g (\mathrm{id} \times S \times S)_h \phi_1 \otimes \phi_2 \otimes f_3 \, d\lambda_x. \tag{16}$$

Averaging over $\Phi \times \Psi$, Theorem 3.1 implies that

$$F := \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} (T \times \mathrm{id} \times T)_g (\mathrm{id} \times S \times S)_h \phi_1 \otimes \phi_2 \otimes f_3$$

exists in $L^2(\lambda)$ and is in fact the orthogonal projection of $\phi_1 \otimes \phi_2 \otimes f_3$ on the space of $T \times \mathrm{id} \times T$- and $\mathrm{id} \times S \times S$-invariant functions. By Observation 4.5 we conclude that the average of $u_{g,h}$ converges, and

$$\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} u_{g,h}(x) = \int K \cdot F \, d\lambda_x \quad (\text{in } L^2(\mu))$$

This concludes the proof of part (1) of Theorem 1.4. □

In order to identify the limit in (3) and prove a combinatorial corollary, we first show that limits over products of Følner sequences can be evaluated by taking iterated limits in each Følner sequence separately.

Claim. Let $\{u_{g,h}\}_{g,h\in G}$ be a bounded sequence indexed by $G \times G$ such that for all two-sided Følner sequences Φ, Ψ, the limits

$$L = \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} u_{g,h}$$

$$L' = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \lim_{m\to\infty} \frac{1}{|\Psi_m|} \sum_{h\in\Psi_m} u_{g,h}$$

exist and are independent of the choice of Følner sequences Φ, Ψ. Then $L = L'$.

Proof. Let Φ, Ψ be two-sided Følner sequences in G. For each n, choose m_n so that for all $g \in \Phi_n$, $\|\frac{1}{|\Psi_{m_n}|}\sum_{h\in\Psi_{m_n}} u_{g,h} - \lim_{m\to\infty} \frac{1}{|\Psi_m|}\sum_{h\in\Psi_m} u_{g,h}\| < \frac{1}{n}$. Then

$$\begin{aligned} L - L' &= \lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_{m_n}|} \sum_{(g,h)\in\Phi_n\times\Psi_{m_n}} u_{g,h} - \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \lim_{m\to\infty} \frac{1}{|\Psi_m|} \sum_{h\in\Psi_m} u_{g,h} \\ &= \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \left(\frac{1}{|\Psi_{m_n}|} \sum_{h\in\Psi_{m_n}} u_{g,h} - \lim_{m\to\infty} \frac{1}{|\Psi_m|} \sum_{h\in\Psi_m} u_{g,h} \right) \\ &= 0 \end{aligned}$$

This proves the claim. □

To prove the inequality claimed in Corollary 1.5, we will apply a corollary of the ergodic theorem for amenable groups. Khintchine ([16]) used the following result in the case where $G = \mathbb{Z}$ to prove a strengthening of the Poincaré recurrence theorem. The argument we give is due to Hopf ([14]).

Corollary 4.6. *Let $\mathbf{X}, G, T$ and Φ be as in the statement of Theorem 3.1, and let $f \in L^2(\mu)$ be non-negative. Then*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \int f T_g \, d\mu \geq \left(\int f \, d\mu \right)^2.$$

Proof. By the ergodic theorem for amenable groups, we have

$$\begin{aligned}
\lim_{n\to\infty} \sum_{g\in\Phi_n} \int f T_g f \, d\mu &= \int f \mathbb{E}(f|\mathcal{I}_T) \, d\mu \\
&= \int \mathbb{E}(f|\mathcal{I}_T)^2 \, d\mu \\
&\geq \left(\int \mathbb{E}(f|\mathcal{I}_T) \, d\mu \right)^2 \\
&= \left(\int f \, d\mu \right)^2,
\end{aligned}$$

where the inequality is an application of Cauchy-Schwarz. □

Proof of Corollary 1.5. Write L for the limit in (4). By Theorem 1.4, the limit L exists, and by the preceding discussion, we can evaluate as follows:

$$\begin{aligned}
L &= \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \lim_{m\to\infty} \frac{1}{|\Psi_m|} \sum_{h\in\Psi_m} \int f T_g f S_h (f T_g f) \, d\mu \\
&\geq \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \left(\int f T_g f \, d\mu \right)^2 \\
&\geq \left(\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \int f T_g f \, d\mu \right)^2 \\
&\geq \left(\left(\int f \, d\mu \right)^2 \right)^2,
\end{aligned}$$

where we have applied either Corollary 4.6 or Cauchy-Schwarz in every step. □

Let $A \subseteq X$ with $\mu(A) > 0$. We can apply Corollary 1.5 with $f = 1_A$ to conclude that intersections $(A \cap T_g^{-1} A) \cap S_h^{-1}(A \cap T_g^{-1} A)$ will have large measure for many pairs (g, h). For this and later corollaries involving syndeticity, we need a general implication relating syndeticity and Følner sequences.

Lemma 4.7. *Let G be an amenable group, and $S \subseteq G \times G$. Then S is both left and right syndetic if and only if for all two-sided Følner sequences Φ, Ψ in G, there exists $n, m \in \mathbb{N}$ such that $S \cap (\Phi_n \times \Psi_m) \neq \emptyset$.*

Proof. Suppose that S is not right syndetic. Let Φ be a Følner sequence in G. Then for all n, m there exists $(x_n, y_m) \in G \times G \setminus \bigcup_{g\in\Phi_n\times\Phi_m} S g^{-1}$, so $(x_n \Phi_n \times y_m \Phi_m) \cap S = \emptyset$ for all n, m, which would contradict an assumption that S meets $x_n \Phi_n \times y_m \Phi_m$ for some n, m.

Now suppose that $(\Phi_n \times \Psi_m) \cap S = \emptyset$ for all n, m. Then $\Phi \times \Psi$ is a Følner sequence in $G \times G$ which does not meet S, so S is not syndetic. □

Corollary 4.8. *Let $(X, \mathcal{B}, \mu), T$ and S be as in Theorem 1.4, and let $A \in \mathcal{B}$. For all $\varepsilon > 0$ the set*

$$R_\varepsilon := \{(g,h) \in G \times G : \mu((A \cap T_g^{-1}A) \cap S_h^{-1}(A \cap T_g^{-1}A)) > \mu(A)^4 - \varepsilon\}$$

is both left- and right syndetic.

Proof. Let $\varepsilon > 0$, and let Φ, Ψ be two-sided Følner sequences in G. Take $f = 1_A$, so that $\mu((A \cap T_g^{-1}A) \cap S_h^{-1}(A \cap T_g^{-1}A)) = \int f T_g f S_h f T_g S_h f \, d\mu$. By Corollary 1.5, there exists n so that

$$\frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h) \in \Phi_n \times \Psi_n} \mu(A \cap T_g^{-1}A \cap S_h^{-1}(A \cap T_g^{-1}A)) > \mu(A)^4 - \varepsilon.$$

Since Φ and Ψ are arbitrary two-sided Følner sequences, this shows that R_ε is syndetic. $\square$

4.3. **Identifying the limit; proof of part (2) of Theorem 1.4.** Evaluating the iterated limit corresponding to (3) leads to an explicit description of the limit. Recall that if $f \in L^\infty(\mu), \mathbb{E}(f|\mathcal{I}_S)(x) = \int f \, d\mu_{\sigma(x)}$ for μ-almost every x. For $f_i \in L^\infty(\mu)$ we have

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{|\Phi_n||\Psi_n|} \sum_{(g,h)\in\Phi_n\times\Psi_n} T_g f_1 S_h(f_2 T_g f_3)(x) \\
&\qquad = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \lim_{m\to\infty} \frac{1}{|\Psi_m|} \sum_{h\in\Psi_m} T_g f_1 S_h(f_2 T_g f_3)(x) \\
&\qquad = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} T_g f_1(x) \int f_2 T_g f_3 \, d\mu_{\sigma(x)} \\
&\qquad = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \int f_2(z) T_g \times T_g f_1 \otimes f_3(x,z) \, d\mu_{\sigma(x)}(z) \\
&\qquad = \int f_2(z) H(x,z) \, d\mu_{\sigma(x)},
\end{aligned}$$

where H is the orthogonal projection of $f_1 \otimes f_3$ on the space of $T \times T$-invariant functions in $L^2(\mu \underset{\mathcal{I}_S}{\times} \mu)$. By symmetry, we find that the limit L is also equal to $\int f_1(z) K(x,z) \, d\mu_{\tau(x)}$, where K is the orthogonal projection of $f_2 \otimes f_3$ on the space of $S \times S$-invariant functions in $L^2(\mu \underset{\mathcal{I}_T}{\times} \mu)$. Recalling the definition of $W_{T/S}$ and $W_{S/T}$, and noting that functions of the form $f \otimes g$ for bounded f, g, span a dense subset of $L^2(\eta)$ whenever η is a measure on $(X \times X, \mathcal{B} \otimes \mathcal{B})$, we conclude that the limits in (3) span a dense subset of $W_{T/S}$ and of $W_{S/T}$. It follows that $W_{T/S} = W_{S/T}$.

The fact that the limits in (3) span $W_{T/S}$ lets us conclude that limits will be constant almost everywhere exactly when the space $W_{T/S}$ consists solely of constant functions. From the definition of $W_{T/S}$ and $W_{S/T}$, we see that this condition is equivalent to the condition that both $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \times \mu, S \times S)$ and $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \times \mu, T \times T)$ are ergodic systems - that is, that both $S \times S$ and $T \times T$ are weakly mixing.

Remarks. (1) This description of the limit allows us to conclude that the factor determined by $W_{T/S}$ is the smallest characteristic factor for the scheme $(T_g, S_h, T_g S_h)$, in the sense that $W_{T/S}$ is a factor of every characteristic factor for the scheme $(T_g, S_h, T_g S_h)$.

(2) The use of the measures λ_x and λ is essentially a "diagonal measures" argument, as in the proof of Corollary 2.4 of [9]. cf. [1].

(3) The method of proof of Theorem 1.4 may be used to prove the following similar result, which appears as Theorem 4.8 in [4].

Theorem 4.9. *Suppose that G is a countable amenable group and Φ is a left Følner sequence for G. Suppose that $(X, \mathcal{B}, \mu)$ is a probability space and that T and S are commuting measure preserving G-actions on X. Then for any $\varphi, \psi \in L^2(X, \mathcal{B}, \mu)$,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \phi(T_g x)\psi(S_g T_g x)$$

exists in $L^1(X, \mathcal{B}, \mu)$.

5. Combinatorial Corollaries

We derive combinatorial corollaries analogous to those in sections 6 and 7 of [4].

Let $\Omega = \{0, 1\}^{G\times G}$ with the product topology, so that Ω is a compact metric space. Regard elements of Ω as functions 1_E, where $E \subseteq G \times G$. Define commuting G-actions T and S on Ω by $(T_g\xi)(g_1, g_2) = \xi(g_1 g, g_2)$ and $(S_g\xi)(g_1, g_2) = (g_1, g_2 g)$.

The following appears as Proposition 6.2 in [4].

Proposition 5.1. *Suppose that G is a countable amenable group and Φ is a left Følner sequence for $G \times G$. Suppose $S \subseteq G \times G$. Let $X = \overline{\{T_g S_h 1_E : g, h \in G\}}$ be the orbit closure of 1_E in $\{0, 1\}^{G\times G}$. If*

$$\bar{d}(E) = \limsup_{n\to\infty} \frac{|E \cap \Phi_n|}{|\Phi_n|} > 0,$$

then there exists a $\{T_g\}$- and $\{S_g\}$-invariant probability measure μ on X such that

$$\mu(\{\eta \in X : \eta(e, e) = 1\}) = \bar{d}(E).$$

(The conclusion stated in [4] was that $\mu(\{\eta \in X : \eta(e, e) = 1\}) > 0$, but the above statement was in fact proved.)

The system constructed in Proposition 5.1 allows us to obtain Corollary 1.6, which is analogous to Theorem 6.1 in [4]. To prove Corollary 1.6, let $E \subseteq G\times G$, and take $(X, \mathcal{B}, \mu, T)$ to be the system constructed from 1_E in Proposition 5.1. The conclusion now follows from Corollary 4.8.

Corollary 5.2. *Let $E \subseteq G \times G$, and let Φ be a left Følner sequence in $G \times G$ with $\limsup_{n\to\infty} \frac{|E\cap\Phi_n|}{|\Phi_n|} = \delta > 0$. Then for all $\varepsilon > 0$, the set*

$$\left\{(g, h) : \limsup_{n\to\infty} \frac{|E \cap E(g, 1_G) \cap E(1_G, h) \cap E(g, h) \cap \Phi_n|}{|\Phi_n|} > \delta^4 - \varepsilon\right\}$$

is syndetic in $G \times G$.

While most of the combinatorial corollaries of recurrence results in measure preserving dynamics pertain to configurations in sets of positive density in groups, Corollary 7.2 of [4] derives the following partition result from the recurrence statement in Theorem 5.2 of [4].

Theorem 5.3. ([4], Corollary 7.2) *Suppose that G is a countable amenable group, $r \in \mathbb{N}$, and that $G \times G \times G = \bigcup_{i=1}^{r} C_i$. Then the set*

$$\{g \in G : \text{ there exists } i, 1 \leq i \leq r, \text{ and } (a,b,c) \in G \times G \times G, \text{ such that } \{(a,b,c),(ga,b,c),(ga,gb,c),(ga,gb,gc)\} \subseteq C_i\}$$

is both left and right syndetic in G.

We derive an analogous result, concerning parallelepiped-like configurations in $G \times G \times G$.

Corollary 5.4. *Suppose that G is a countable amenable group, $r \in \mathbb{N}$, and that $G^3 = \bigcup_{i=1}^{r} C_i$. For $a = (a_1, a_2, a_3) \in G^3$, let*

$$(g,h,k) \cdot (a_1,a_2,a_3) = \{(a_1,a_2,a_3),(ga_1,a_2,a_3),(a_1,ha_2,a_3),(a_1,a_2,ka_3),(ga_1,ha_2,a_3),(ga_1,a_2,ka_3),(a_1,ha_2,ka_3),(ga_1,ha_2,ka_3)\}.$$

Then the set

$$\{(g,h,k) \in G \times G \times G : \text{ there exists } i, 1 \leq i \leq r, \text{ and } a \in G^3 \text{such that } (g,h,k) \cdot a \subseteq C_i\}$$

is syndetic in $G \times G \times G$.

We will derive Corollary 5.4 from the following topological recurrence statement.

Theorem 5.5. *Let R, S and T be commuting actions of G by homeomorphisms of a compact metric space (X,d). Then for all $\varepsilon > 0$, the set*

$$J_\varepsilon = \{(g,h,k) \in G \times G \times G : \text{ there exists } x \in X \text{ such that } \operatorname{diam}\{x, T_g x, S_h x, R_k x, T_g S_h x, T_g R_k x, S_h R_k x, T_g S_h R_k x\} < \varepsilon\}$$

is both left and right syndetic in $G \times G \times G$.

In the proof of Theorem 5.5 we will use the following fact: Given minimal topological system, (X,d), with acting group Γ, a point $x \in X$ and an open set $U \subset X$, the set $\{\gamma \in \Gamma : \gamma x \in U\}$ is syndetic in Γ. If Γ is amenable, combining this fact with the ergodic theorem for amenable groups we see that every Borel probability measure on X which is preserved by the action of Γ assigns positive measure to every open set.

We remark that of the consequences of Corollary 4.6 is that given a system $(X, \mathcal{B}, \mu, T)$, with T a measure preserving action of an amenable group, and a set $A \in \mathcal{B}$ of positive measure, the set $\{g \in G : \mu(A \cap T_g^{-1} A) > 0\}$ is both left- and right syndetic.

Proof. We may assume that the group generated by the individual elements of R, S, and T acts minimally on X.

Claim. For all open $U \subseteq X$, there exists a syndetic set $H \subseteq G \times G$ such that for all $(g,h) \in H$, there exists $y \in U$ such that $\{y, T_g y, S_h y, T_g S_h y\} \subseteq U$.

To prove the claim, let $Y \subseteq X$ be a closed subset of X minimal with respect to the group generated by the individual elements of S and T. Let ν be a measure on Y invariant with respect to S and T. By minimality, we have, for all open U with $U \cap Y \neq \emptyset$, $\nu(U) > 0$. Fixing such a U, we let H be the set of (g,h) such that $\nu(U \cap T_g^{-1} U \cap S_h^{-1} U \cap T_g^{-1} S_h^{-1} U) > 0$, so that H is syndetic, by Corollary 4.8. Now suppose that U is an arbitrary open subset of X. By minimality of the action generated by the elements of R, S and T, there exists $k \in G$ so that $Y \cap R_k^{-1} U \neq \emptyset$, so by the above argument the set H of (g,h) such that

$\nu(R_k^{-1}U \cap T_g^{-1}R_k^{-1}U \cap S_h^{-1}R_k^{-1}U \cap T_g^{-1}S_h^{-1}R_k^{-1}U) > 0$ is syndetic, and so for all $(g,h) \in H, U \cap T_g^{-1}U \cap S_h^{-1}U \cap T_g^{-1}S_h^{-1}U \neq \emptyset$. Thus, H satisfies the conclusion of the claim.

Now let $\varepsilon > 0$. Let Φ, Ψ, and Θ be Følner sequences in G. Let U be an open set in X of diameter less than ε, and let $H \subseteq G \times G$ be a syndetic set such that for all $(g,h) \in H$, there exists $y \in U$ with $\{y, T_g y, S_h y, T_g S_h y\} \subseteq U$. Since H is syndetic, there exists $n \in \mathbb{N}$ such that $H \cap (\Phi_n \times \Psi_n) \neq \emptyset$. Fix such $(g,h) \in H \cap (\Phi_n \times \Psi_n)$ and $y \in U$. Let V be a neighborhood of y such that $\{x, T_g x, S_h x, T_g S_h x\} \subseteq U$ for all $x \in V$. By minimality of the action generated by R, S, and T, there is an R,S,T-invariant measure μ on X with $\mu(V) > 0$. By Poincaré recurrence, $K := \{k \in G : R_k^{-1}V \cap V \neq \emptyset\}$ is syndetic, so for all large enough m, there exists $k \in K \cap \Theta_m$. For such $k, x \in R_k^{-1}V \cap V$, we have

$$\{x, T_g x, S_h x, T_g S_h x, R_k x, T_g R_k x, S_h R_k x, T_g S_h R_k x\} \subseteq U.$$

This shows that J_ε meets $\Phi_n \times \Psi_n \times \Theta_m$. Since Φ, Ψ, and Θ are arbitrary Følner sequences, Lemma 4.7 implies that J_ε is both left and right syndetic. $\square$

To derive Corollary 5.4 from Theorem 5.5, we apply a standard construction of a topological dynamical system from the partition $G \times G \times G = \bigcup_{i=1}^r C_i$ in the hypothesis of Corollary 5.4. Let $\Lambda = \{1, \ldots, r\}$, and consider the shift space $X = (G \times G \times G)^\Lambda$, with the product topology, together with the actions R, S, T of G defined by $(T_g\xi)x,y,z) = \xi(g^{-1}x, y, z), (S_g\xi)x,y,z) = \xi(x, g^{-1}y, z), (R_g\xi)x,y,z) = \xi(x,y,g^{-1}z)$. Define a point ξ_0 in X by $\xi(x,y,z) = i$ if $(x,y,z) \in C_i$. Let Y be the orbit closure of x_0 under the group generated by R, S and T. By Theorem 5.5, the set of $(g,h,k) \in G^3$ such that there exists $\xi \in Y$ with $\operatorname{diam}\{\xi, T_g\xi, S_h\xi, T_gS_h\xi, R_k\xi, T_gR_k\xi, S_hR_k\xi, T_gS_hR_k\xi\} < 1$ is both left and right syndetic. In particular, the set of such (g,h,k) such that there exists ξ such that every element of $\{\xi, T_g\xi, S_h\xi, T_gS_h\xi, R_k\xi, T_gR_k\xi, S_hR_k\xi, T_gS_hR_k\xi\}$ agrees at the (e,e,e) coordianate is both left and right syndetic. Taking inverses gives the conclusion of Corollary 5.4.

6. An example.

In this section we specialize to the case $G = \mathbb{Z}$, and give an example of a class of systems where the characteristic factor for the scheme (T_g, S_h, T_gS_h) is not an inverse limit of nilsystems. This is in contrast to the case $S = T$, where the characteristic factor (assuming ergodicity of T) is a compact group rotation, as Bergelson showed in [3]. We thank Alexander Leibman for suggesting this example. (cf. the example presented in Section 3 of [8] .)

Let $\mathbf{Y}_0 = (Y_0, \mathcal{D}_0, \mu_0, T_0), \mathbf{Y}_1 = (Y_1, \mathcal{D}_1, \mu_1, S_1)$ be ergodic $\mathbb{Z}$-systems, and let K be a compact metric group with Haar measure m. Let $\tau : Y_0 \to K, \sigma : Y_1 \to K$ be measurable maps. Define transformations T and S on $X := Y_0 \times Y_1 \times K$ by $T(y_0, y_1, k) = (T_0y_0, y_1, \tau(y_0)k), S(y_0, y_1, k) = (y_0, S_1y_1, k\sigma(y_1)^{-1})$. Then T and S commute, and preserve $\mu := \mu_0 \times \mu_1 \times m$. If τ is chosen so that $(y_0, k) \mapsto (T_0y_0, \tau(y_0)k)$ is ergodic, then the group generated by S and T acts ergodically on X. One can verify that $W_{T/S} = L^2(\mu)$ in this situation, while such systems $(X, \mathcal{B}, \mu, T, S)$ are usually not isomorphic to inverse limits of nilsystems. This is in sharp contrast to the characteristic factors for schemes involving powers of a single transformation, such as $(T^n, T^m, T^r, T^{n+m}, T^{n+r}, T^{m+r}, T^{n+m+r})$, which are shown in [15] to be inverse limits of nilsystems.

References

[1] T. Austin. On the norm convergence of nonconventional ergodic averages. Preprint. arXiv:0805.0320

[2] V. Bergelson. Ergodic Ramsey Theory. *Logic and combinatorics*, 63-87, Contemp. Math., **65**, Amer. Math. Soc., Providence, RI, 1987.

[3] V. Bergelson. The multifarious Poincaré recurrence theorem. *Descriptive set theory and dynamical systems LMS Lecture Note Series 277*, Cambridge Univ. Press, Cambridge, 2000 pp. 31-57.

[4] V. Bergelson, R. McCutcheon, and Q. Zhang. A Roth theorem for amenable groups. *Amer. J. Math.* **119** (1997), 1173-1211.

[5] V. Bergelson and J. Rosenblatt. Joint ergodicity for group actions. *Ergod. Th. & Dynam. Sys.* **8** (1988), 351-364.

[6] J-P. Conze and E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112** (1984), 143-175.

[7] J-P. Conze and E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *Probabilités*, 1-31, Publ. Inst. Rech. Math. Rennes, 1987-1, Univ. Rennes I, Rennes, 1988.

[8] N. Frantzikinakis and B. Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dyn. Systems* **25** (2005), 799-809.

[9] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204-256.

[10] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, NJ, 1981.

[11] H. Furstenberg. Nonconventional ergodic averages. *The legacy of John von Neumann. Proc. Sympos. Pure Math. 50*, Amer. Math. Soc., Providence, RI, 1990. pp. 43-56.

[12] H. Furstenberg and B. Weiss. A mean ergodic theorem for $(1/N)\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$. *Convergence in ergodic theory and probability. Ohio State Univ. Math. Res. Inst. Publ. 5.* de Gruyter, Berlin, 1996. pp.193-227.

[13] F. P. Greenleaf. Invariant means on topological groups and their applications. *Van Nostrand Mathematical Studies 16*, Van Nostrand, 1969.

[14] E. Hopf. *Ergodentheorie.* Chelsea, New York. 1948.

[15] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.* **161** (2005), 397-488.

[16] A. Khintchine. Eine Verschärfung des Poincaréschen "Wiederkehrsatzes". *Compositio Math.* **1** (1935), 177-179.

[17] H.L. Royden. *Real Analysis*, third edition, Prentice Hall, 1988.

[18] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems.* **28** (2008), 657-688.

[19] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), 53-97.

[20] Q. Zhang. On convergence of the averages $(1/N)\sum_{n=1}^{N} f_1(R^n x)f_2(S^n x)f_3(T^n x)$. *Monatsh. Math.* **122** (1996), 275-300.

Department of Mathematics, The Ohio State University, 231 W. 18th Ave., Columbus, OH 43212

E-mail address: griesmer@math.osu.edu